AD-A073 577 PURDUE UNIV LAFAYETTE IND DEPT OF STATISTICS F/G 12/1

ASYMPTOTIC THEORY FOR PROCESS LEAST SQUARES ESTIMATORS FOR DIFF--ETC(U)

JUL 79 B L RAO, H RUBIN N00014-75-C-0455

UNCLASSIFIED MMS-79-13 NL

1 OF 1

AD
A073577

END
DATE
FILMED
10-79
DDC

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

(12) LEVEL II

(6)

ASYMPTOTIC THEORY FOR PROCESS LEAST SQUARES
ESTIMATORS FOR DIFFUSION PROCESSES.

by

(10)
B.L.S. Prakasa Rao* and Herman Rubin**
Indian Statistical Institute, New Delhi
and
Purdue University

Department of Statistics
Division of Mathematical Sciences
(9) Mimeograph Series # 79-13 (14) MMS-

(11) July 1979

(12) 29p.

(15) * Research partially supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University.
NSF-MCS76-08316
** Research partially supported by NSF Grant MCS76-08316.

291 730

1B

ABSTRACT

Strong consistency and asymptotic normality of an estimator related to least squares estimator for parameters involved in nonlinear stochastic differential equations are investigated by studying families of stochastic integrals using Fourier analytic methods.

AMS (1980) Subject Classification: Primary 62M05, Secondary 60H10

Key words and phrases: Stochastic Differential Equation; Diffusion Process; Least Squares Estimation; Consistency; Asymptotic Normality

1. Introduction

Recently there is a growing interest in the study of inference problems for stochastic processes both continuous and discrete time in view of the large number of applications to engineering problems. It has been found that the class of diffusion processes is amenable for statistical analysis. A survey of the recent work in this area is given in Basawa and Prakasa Rao (1979). Further work on asymptotic theory of maximum likelihood and Bayes estimators for parameters of diffusion processes is discussed in Prakasa Rao (1979a).

Dorogovchev (1976) studied weak consistency of least square estimators for parameters of diffusion processes which are solutions of non-linear stochastic differential equations. Asymptotic normality and asymptotic effiency of these estimators is investigated in Prakasa Rao (1979b). Our aim in this paper is to study limiting properties of a process related to least squares estimator and hence to discuss the asymptotic properties of an estimator derived from the limiting process. We study strong consistency and asymptotic normality of this estimator. Our approach here is entirely different from that of Dorogovchev (1976) and Prakasa Rao (1979b). We believe that our techniques for study of families of stochastic integrals is new and is of independent interest.

2. Study of process related to least squares estimator

Let $\{X(t), t \geq 0\}$ be a real-valued stationary ergodic process satisfying the stochastic differential equation

$$dX(t) = f(\theta_0, X(t))dt + d\xi(t), \quad X(0) = X_0, \quad t \geq 0$$

where $\xi(t)$ is a Wiener process with mean zero and variance $\sigma^2 t$, σ^2 being unknown and $E[X_0^2] < \infty$. Suppose $f(\theta, x)$ is a known real-valued function continuous on $\Theta \times R$ where Θ is a closed interval on the real line and $\theta_0 \in \Theta$ is unknown. Without loss of generality, assume that $\Theta = [-1,1]$.

Suppose the process $\{X(t), 0 \leq t \leq T\}$ is observed at time points t_k, $k = 0,1,\ldots,n-1$ with $t_0 = 0$ and $t_n = T$. Let

$$Q_n^T(\theta) = \sum_{k=0}^{n-1} \frac{[X(t_{k+1}) - X(t_k) - f(\theta, X(t_k))\Delta t_k]^2}{t_k} .$$

where $\Delta t_k = t_{k+1} - t_k$, $\theta \leq k \leq n-1$. An estimator $\hat{\theta}_{n,T}$ which minimizes $Q_n^T(\theta)$ over $\theta \in \Theta$ is called a <u>least squares estimator</u> of θ. Assume that such an estimator exists. Note that if $\hat{\theta}_{n,T}$ minimizes $Q_n^T(\theta)$, then it minimizes $Q_n^T(\theta) - Q_n^T(\theta_0)$.

We shall first study the limiting properties of the process $\{Q_n^T(\theta) - Q_n^T(\theta_0),\ \theta \in \Theta\}$ as the norm of division $\Delta_n = \max_{1 \leq k \leq n} |t_{k+1} - t_k|$ tends to zero. Let $\Delta X_k = X(t_{k+1}) - X(t_k)$ and $\Delta\xi_k = \xi(t_{k+1}) - \xi(t_k)$, $0 \leq k \leq n-1$. For any fixed θ,

$$Q_n^T(\theta)-Q_n^T(\theta_0)$$

$$= \sum_k \frac{1}{\Delta t_k} [\Delta X_k - f(\theta, X(t_k))\Delta t_k]^2$$

$$- \sum_k \frac{1}{\Delta t_k} [\Delta X_k - f(\theta_0, X(t_k))\Delta t_k]^2$$

$$= \sum_k \frac{1}{\Delta t_k} \{\int_{t_{k-1}}^{t_{k+1}} f(\theta_0, X(t))dt + \Delta\xi_k - f(\theta, X(t_k))\Delta t_k\}^2$$

$$- \sum_k \frac{1}{\Delta t_k} \{\int_{t_k}^{t_{k+1}} f(\theta_0, X(t))dt + \Delta\xi_k - f(\theta_0, X(t_k))\Delta t_k\}^2$$

$$= \sum_k \frac{1}{\Delta t_k} \{\int_{t_k}^{t_{k+1}} [f(\theta_0, X(t)) - f(\theta, X(t_k))]dt + \Delta\xi_k\}^2$$

$$- \sum_k \frac{1}{\Delta t_k} \{\int_{t_k}^{t_{k+1}} [f(\theta_0, X(t)) - f(\theta_0, X(t_k))]dt + \Delta\xi_k\}^2 .$$

It is easy to check that

(2.0) $Q_n^T(\theta)-Q_n^T(\theta_0)$

$$= \sum_k [f(\theta_0, X(t_k)) - f(\theta, X(t_k))]^2 \Delta t_k$$

$$+ 2 \sum_k [f(\theta_0, X(t_k) - f(\theta, X(t_k)]\Delta\xi_k$$

$$+ 2 \sum_k \{f(\theta_0, X(t_k) - f(\theta, X(t_k))\} \int_{t_k}^{t_{k+1}} \{f(\theta_0, X(t)) - f(\theta_0, X(t_k))\}dt$$

$$= I_{1n} + 2I_{2n} + 2I_{3n} .$$

Assume that the regularity condition on $f(x,\theta)$ stated at the end of this section are satisfied. Since $f(\theta,x)$ is continuous in x and the

process X has continuous sample paths with probability one, it follows that

$$(2.1) \qquad I_{1n} \xrightarrow{a.s.} \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]^2 dt$$

as $\Delta_n \to 0$. Assumption (A2) implies that

$$(2.2) \qquad I_{2n} \xrightarrow{q.m.} \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]d\xi(t)$$

as $\Delta_n \to 0$ in view of stationarity of the process X where the last integral is the Ito-stochastic integral.

Let us now estimate I_{3n}. In view of assumption (A4), it can be checked that

$$(2.3) \qquad \left|\int_{t_k}^{t_{k+1}} \{f(\theta_0,X(t))-f(\theta_0,X(t_k))\}dt\right|$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} |X(t)-X(t_k)|dt$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} \{|\xi(t)-\xi(t_k)|+\int_{t_k}^{t} f(\theta_0,X(s))|ds\}dt$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} |\xi(t)-\xi(t_k)|dt+L^2(\theta_0)\int_{t_k}^{t_{k+1}} \left[\int_{t_k}^{t}\{1+|X(s)|\}ds\right] dt$$

$$\leq L(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} |\xi(t)-\xi(t_k)|+L^2(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} \int_{t_k}^{t} \{1+|X(s)|\}ds.$$

$$\leq L(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} |\xi(t)-\xi(t_k)|+L^2(\theta_0)\Delta t_k^2 \sup_{t_k\leq t\leq t_{k+1}} \{1+|X(t)|\}$$

for $0 \leq k \leq n-1$. Using assumption (A4) again, we obtain the following inequality:

$$(2.4)\qquad I_{3n} \le \sum_k J(X(t_k))\Big\{L(\theta_0)\cdot\Delta t_k \sup_{t_k\le t\le t_{k+1}} |\xi(t)-\xi(t_k)| + L^2(\theta_0)\Delta t_k^2 \sup_{t_k\le t\le t_{k+1}} \{1+|X(t)|\}\Big\}|\theta-\theta_0|.$$

Since $J(\cdot)$ is continuous and $X(\cdot)$ has continuous sample paths almost surely, it follows that there exists a constant $C^*(\theta_0)$ depending on T only such that

$$(2.5)\qquad I_{3n} \le C^*(\theta_0)\Big\{\sum_k \Delta t_k\cdot \sup_{t_k\le t\le t_{k+1}} |\xi(t)-\xi(t_k)| + \sum_k \Delta t_k^2\Big\}|\theta-\theta_0| .$$

Since $\theta\in\Theta$ compact, it follows that

$$I_{3n} \le C(\theta_0)\{\sum_k \Delta t_k(1+\Delta t_k)(2\Delta t_k \log 1/\Delta t_k)^{1/2} + \sum_k \Delta t_k^2\} \quad \text{a.s.}$$

whenever Δ_n is sufficiently small by the law of iterated logarithm for Brownian increments (cf. McKean (1969), p.14). Therefore

$$(2.6)\qquad I_{3n} = O(\sum_k \Delta t_k^{3/2} \log^{1/2} 1/\Delta t_k) \text{ a.s.}$$

uniformly in $\theta\in\Theta$. Furthermore the convergence in (2.1) is uniform in $\theta\in\Theta$ since

$$|f(\theta_0,X(t))-f(\theta,X(t))|^2 \le |\theta_0-\theta|^2 J^2(X(t)) \le C\, J^2(X(t))$$

and $J(X(t))$ is integrable pathwise on $[0,T]$ by (A4). Here we have used the fact that Θ is compact. Hence

$$(2.7)\qquad I_{1n} = \int_0^T [f(\theta_0,X(t)-f(\theta,X(t)]^2 dt + o(1) \quad \text{a.s.}$$

uniformly in θ as $\Delta_n \to 0$. We shall discuss uniform convergence of I_{2n} in the next section.

Relations (2.0), (2.6) and (2.7) show that, for any fixed T,

$$(2.8) \qquad Q_n^T(\theta)-Q_n^T(\theta_0) = \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]^2dt + I_{2n}+o(1) \text{ a.s.}$$

uniformly in $\theta \in \Theta$ compact as $\Delta_n \to 0$ where I_{2n} satisfies relation (2.2).

Let us consider the limiting process

$$(2.9) \qquad R_T(\theta) = \int_0^T [f(\theta_0,X(t)) - f(\theta,X(t))]^2dt$$

$$+ 2\int_0^T [f(\theta_0,X(t)) - f(\theta,X(t))]d\xi(t)$$

$$= \int_0^T v^2(\theta,X(t))dt - 2\int_0^T v(\theta,X(t))d\xi(t)$$

where

$$(2.10) \qquad v(\theta,x) = f(\theta,x) - f(\theta_0,x).$$

We study the limiting properties of the process $\{R_T(\theta), \theta \in \Theta\}$ in the next section.

<u>Assumptions</u>

(A1) $f(\theta,x)$ is continuous in (θ,x) and differentiable with respect to θ. Denote the first partial derivative of f with respect to θ by $f_\theta^{(1)}(\theta.x)$ and the derivative evaluated at θ_0 by $f_\theta^{(1)}(\theta_0,x)$.

(A2) $E[f_\theta^{(1)}(\theta_0,X(0))]^2 < \infty$

(A3) $f_\theta^{(1)}(\theta,x)$ is Lipschitzian in θ for each x i.e., there exists $\alpha > 0$ such that

$$|f_\theta^{(1)}(\theta,x)-f_\theta^{(1)}(\phi,x)| \le c(x)|\theta-\phi|^\alpha , \quad x \in R,\theta,\phi \in \Theta$$

and

$$E[c^2(X(0))] < \infty.$$

(A4) $f(\theta,x)$ satisfies the following conditions:

(i) $|f(\theta,x)| \le L(\theta)(1+|x|)$, $\theta \in \Theta$, $x \in R$; $\sup\{L(\theta): \theta \in \Theta\} < \infty$.

(ii) $|f(\theta,x)-f(\theta,y)| \le L(\theta)|x-y|$, $\theta \in \Theta$, $x,y, \in R$.

(iii) $|f(\theta,x)-f(\phi,x)| \le J(x)|\theta-\phi|$, $\theta , \phi \in \Theta$, $x \in R$

where $J(\cdot)$ is continuous and $E[J^2(X(0))] < \infty$.

(A5) $I(\theta) \equiv E[f(\theta,X(0))-f(\theta_0,X(0)]^2 > 0$ for $\theta \neq \theta_0$.

<u>Remark</u>: Since $E[X^2(0)] < \infty$, assumption A4(i) implies that

$$E[f(\theta,X(0))]^2 < \infty$$

for all $\theta \in \Theta$.

3. Study of a limiting process related to least squares estimator

Let us now study the properties of the limiting process

$$Z_T(\theta) \equiv \frac{1}{\sqrt{T}} \int_0^T v(\theta,X(t))d\xi(t) \tag{3.1}$$

as a process in the parameter $\theta \in \Theta = [-1,1]$ as $T \to \infty$. From the central limit theorem for stochastic integrals (cf. Basawa and Prakasa Rao (1979)), it can be shown that

$$\frac{1}{\sqrt{T}} \int_0^T v(\theta,X(t))d\xi(t) \xrightarrow{\mathcal{L}} N(0,E[v(\theta,X(0))]^2\sigma^2)$$

since the process X is stationary ergodic. In general, finite dimensional distributions of the process $\{Z_T(\theta), \theta \in \Theta\}$ converge to the finite dimensional distributions of the Gaussian process $\{Z(\theta), \theta \in \Theta\}$ with mean zero and covariance function

$$R(\theta_1,\theta_2) = E[v(\theta_1,X(0))v(\theta_2,X(0))]\sigma^2.$$

We shall now prove the weak convergence of the process $\{Z_T(\theta), \theta \in \Theta\}$ on C[-1,1] under uniform norm. It is sufficient to prove that

$$\lim_{T\to\infty} \overline{\lim_{\delta\to 0}} P\left(\sup_{|\theta-\phi|<\delta} |Z_T(\theta)-Z_T(\phi)| > \varepsilon\right) = 0. \tag{3.2}$$

Since $v(\theta,x)$ is differentiable with respect to θ on $[-1,1]$ by assumption (A1), it is easy to see that there exists a cubic polynomial $g(\theta,x)$ in θ such that

$$g(-1,x) = v(-1,x),\ g(1,x) = v(1,x)$$

and

$$g_\theta^{(1)}(-1,x) = v_\theta^{(1)}(-1,x),\ g_\theta^{(1)}(1,x) = v_\theta^{(1)}(1,x).$$

Let

$$h(\theta,x) = v(\theta,x)-g(\theta,x).$$

Then $h(-1,x) = h(1,x) = 0, h_\theta^{(1)}(-1,x) = h_\theta^{(1)}(1,x) = 0$. Now

$$Z_T(\theta) = \frac{1}{\sqrt{T}} \int_0^T h(\theta,x(t))d\xi(t) + \frac{1}{\sqrt{T}} \int_0^T g(\theta,X(t))d\xi(t). \tag{3.3}$$

Since $g(\theta,x)$ is a cubic polynomial in θ with coefficients in x which are linear functions of $v(-1,x)$, $v(1,x)$, $v_\theta^{(1)}(-1,x)$ and $v_\theta^{(1)}(1,x)$, it is easy to check the uniform equi-continuity condition of type (3.2) for

$$\frac{1}{\sqrt{T}} \int_0^T g(\theta,X(t))d\xi(t).$$

Let us now consider the process

$$W_T(\theta) = \frac{1}{\sqrt{T}} \int_0^T h(\theta,X(t))d\xi(t). \tag{3.4}$$

Let the Fourier expansion for $h(\theta,x)$ in $L_2([-1,1])$ be given by

$$h(\theta,x) = \sum_n a_n(x)e^{\pi i n\theta}, \qquad x \in R. \tag{3.5}$$

<u>Lemma 3.1</u>

$$\int_0^T h(\theta,X(t))d\xi(t) = \sum_n \{ \int_0^T a_n(X(t))d\xi(t)\}e^{\pi i n\theta} \tag{3.6}$$

in the sense of convergence in quadratic mean.

Proof An approximating sum in L_2-norm for

$$\int_0^T h(\theta,X(t))d\xi(t)$$

is

$$A_{1N} = \sum_{j=1}^{N} h(\theta,X(t_{j-1}))\Delta\xi_j$$

and an approximating sum in L_2-norm for $\sum_n \{\int_0^T a_n(X(t))d\xi(t)\}e^{\pi in\theta}$ is

$$A_{2NM} = \sum_{|n|\leq M} e^{\pi in\theta}(\sum_{j=1}^{N} a_n(X(t_{j-1})\Delta\xi_j) .$$

It is sufficient to prove that $E|A_{1N}-A_{2NM}|^2 \to 0$ as $N \to \infty$ and $M \to \infty$. Now

$$E|A_{1N}-A_{2NM}|^2 = E|\sum_{j=1}^{N}\{(h(\theta,X(t_{j-1}))-\sum_{n=-M}^{M} e^{\pi in\theta}a_n(X(t_{j-1}))\}\Delta\xi_j|^2$$

$$= E|\sum_{j=1}^{N}\sum_{|n|>M} a_n(X(t_{j-1}))e^{\pi in\theta}\Delta\xi_j|^2$$

$$\leq [\sum_{|n|>M}\{E(\sum_{j=1}^{N} a_n(X(t_{j-1}))\Delta\xi_j)^2\}^{\frac{1}{2}}]^2$$

by the elementary inequality

$$E|\sum_n \lambda_n Y_n|^2 \leq (\sum_n|\lambda_n|(E(Y_n^2))^{\frac{1}{2}})^2$$

for any sequence of complex numbers $\{\lambda_n\}$ and any sequence of real valued random variables $\{Y_n, n \geq 1\}$. Hence

$$E|A_{1N}-A_{2NM}|^2 \leq [\sum_{|n|>M}\{\sum_{j=1}^{N} E(a_n(X(t_{j-1}))^2\Delta t_j\}^{\frac{1}{2}}]^2 .$$

Since

$$\sum_{j=1}^{N} E(a_n(X(t_{j-1}))^2\Delta t_j \to \int_0^T E\{a_n(X(t)\}^2 dt = T\mu_n \quad \text{(say)},$$

as $N \to \infty$, it is sufficient to prove that $\sum_n \mu_n^{\frac{1}{2}} < \infty$. This follows from remarks following Lemma 3 of the appendix under assumption (A3).

Let

$$(3.7) \qquad W_n = \frac{1}{\sqrt{T}} \int_0^T a_n(X(t))d\xi(t).$$

Lemma 3.2. For every $\varepsilon > 0$,

$$(3.8) \qquad \lim_{\delta\to 0} P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big) = 0$$

for every $T > 0$.

Proof. In view of Lemma 3.1, for any $\varepsilon > 0$,

$$(3.9) \qquad P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big)$$

$$= P\Big(\sup_{|\theta-\phi|<\delta} \Big|\sum_{n=-\infty}^{\infty} W_n(e^{\pi in\theta}-e^{\pi in\phi})\Big| > \varepsilon\Big)$$

$$\leq P\Big(\sup_{|\theta-\phi|<\delta} \sum_{n=-\infty}^{\infty} |W_n|\,|e^{\pi in\theta}-e^{\pi in\phi}| > \varepsilon\Big).$$

Let n_0 be chosen so that

$$(3.10) \qquad \sum_{n=n_0}^{\infty} \mu_n^{1/3} < \varepsilon 2^{-4/3}.$$

This is possible since $\sum_{n=1} \mu_n^{1/3} < \infty$ by Lemma 3 of the appendix.

Inequality (3.9) implies that

$$P\left(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\right)$$

$$\leq P\left(\sup_{|\theta-\phi|<\delta} \sum_{n=-n_0}^{n_0} |W_n|n|\theta-\phi| > \frac{\varepsilon}{4\pi}\right) + P\left(\sum_{|n|>n_0} |W_n| > \frac{\varepsilon}{2}\right)$$

$$\leq \sum_{n=1}^{n_0} P\left(|W_n| > \frac{\varepsilon}{2\pi n_0 \delta}\right) + 2\sum_{n=n_0+1}^{\infty} P(|W_n| > \varepsilon_n)$$

$$\left(\text{Here } \varepsilon_n = \frac{\varepsilon}{2^{4/3}} \mu_n^{1/3} \left(\sum_{n=n_0+1}^{\infty} \mu_n^{1/3}\right)^{-1}\right)$$

$$\leq \left(\frac{2\pi n_0 \delta}{\varepsilon}\right)^2 \sum_{n=1}^{n_0} \mu_n + \sum_{n=n_0+1}^{\infty} \frac{\mu_n}{\varepsilon_n^2}$$

$$(\text{since } E(W_n) = 0 \text{ and } Var(W_n) = \mu_n)$$

$$= \frac{(2\pi n_0 \delta)^2}{\varepsilon^2} \sum_{n=1}^{n_0} \mu_n + \frac{8}{\varepsilon^2}\left(\sum_{n=n_0+1}^{\infty} \mu_n^{1/3}\right)^3$$

$$= C_{n_0} \frac{\delta^2}{\varepsilon^2} + \frac{8}{\varepsilon^2}\left(\frac{\varepsilon}{2}\right)^3$$

where C_{n_0} depends only on n_0. Choosing δ such that

$$C_{n_0} \frac{\delta^2}{\varepsilon^2} < \varepsilon \quad \text{i.e.} \quad 0 < \delta < \left(\frac{\varepsilon^3}{2C_{n_0}}\right)^{\frac{1}{2}}$$

we have the inequality

$$P\left(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\right) \leq 2\varepsilon$$

for every $0 < \delta < \left(\frac{\varepsilon^3}{2Cn_0}\right)^{\frac{1}{2}}$ and for every $T > 0$. This proves (3.8).

Theorem 3.1. The family of stochastic processes $\{Z_T(\theta), \theta \in \Theta\}$ on $C[-1,1]$ converge in distribution to the Gaussian process with mean zero and covariance function

$$R(\theta_1,\theta_2) = E[v(\theta_1,X(0))v(\theta_2,X(0))]\sigma^2$$

as $T \to \infty$.

4. Strong consistency

Let us now consider the limiting process $R_T(\theta)$ defined by (2.9). Any estimator $\hat{\theta}_T$ which minimizes

$$(4.1) \quad R_T(\theta) \equiv \int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt$$

$$-2\int_0^T [f(\theta,X(t))-f(\theta_0,X(t))]d\xi(t)$$

is called a process least squares estimator of θ.

Let μ_θ be the measure generated by the process X on C[0,T] when θ is the true parameter. From the general theory of diffusion processes, the Radon-Nikodym derivative of μ_θ with respect to μ_{θ_0} exists and is given by

$$(4.2) \qquad \frac{d\mu_\theta}{d\mu_{\theta_0}} = \exp\left\{\int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}d\xi(t) - \frac{1}{2}\int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt\right\}.$$

(cf. Gikhman and Skorokhod (1972), p.90). Hence

$$\log \frac{d\mu_\theta}{d\mu_{\theta_0}} = -\frac{1}{2} R_T(\theta)$$

which proves that the process least squares estimator $\hat{\theta}_T$ is the same as the maximum likelihood estimator $\tilde{\theta}_T$ of θ(cf. Basawa and Prakasa Rao (1979)) when the process X is observed over [0,T].

Let

$$(4.3) \qquad I_T(\theta) \equiv \int_0^T [f(\theta,X(t))-f(\theta_0,X(t))]^2 dt$$

and W* be a standard Wiener process. Since the solution of the stochastic differential equation given in Section 2 is stationary ergodic by hypothesis, it follows that $I_T(\theta) \to \infty$ a.s. for $\theta \neq \theta_0$ by (A5) and the process $\{R_T(\theta)\}$ can be identified with the process $\{I_T(\theta) + 2W^*(T_T(\theta))\}$. Furthermore

$$(4.4) \qquad I_T(\theta) + 2W^*(T_T(\theta)) \to \infty \quad \text{a.s.}$$

as $T \to \infty$ *for any* $\theta = \theta_0$. Hence θ and θ_0 are pairwise consistent. Note that

$$(4.5) \qquad R_T(\theta) = I_T(\theta) + \sqrt{T}\, Z_T(\theta), \quad \theta \in \Theta, \quad T \geq 0$$

where $I_T(\theta)$ is defined by (4.3) and $Z_T(\theta)$ is given by (3.1). Let

$$(4.6) \qquad Z_T^*(\theta) = \sqrt{T}\, Z_T(\theta).$$

Then

$$(4.7) \qquad \frac{1}{T} I_T(\theta) \to I(\theta) \quad \text{a.s. as } T \to \infty \text{ by the ergodic theorem.}$$

In order to study the strong consistency of the estimator $\hat{\theta}_T$, we shall first obtain bounds on the modulus of coninuity of $I_T(\theta)$ and $Z_T^*(\theta)$.

<u>Lemma 4.1.</u> Under the assumptions (A1)-(A5),

$$|I_T(\theta)-I_T(\phi)| \leq C_1|\theta-\phi| \int_0^T J(X(t))(1+|X(t)|)dt \quad \text{a.s.}$$

where C_1 is a constant independent of T, θ and ϕ.

<u>Proof</u>. Note that

$$I_T(\theta)-I_T(\phi) = \int_0^T \{f(X(t),\theta)-f(X(t),\phi)\} \cdot$$

$$\cdot \{f(X(t),\phi)+f(X(t),\theta)-2f(X(t),\theta_0)\} dt$$

and therefore

$$|I_T(\theta)-I_T(\phi)| \le |\theta-\phi| \int_0^T J(X(t))\cdot\{L(\theta)+L(\phi)+2L(\theta_0)\}\{1+|X(t)|\}dt$$

$$\le C_1|\theta-\phi| \int_0^T J(X(t)\{1+|X(t)|\}dt \ .$$

Remark. Since $E[J^2(X(0)] < \infty$ and $E[X^2(0)] < \infty$, it follows that $E[J(X(0))X(0)] < \infty$ and hence by the ergodic theorem

$$\frac{1}{T}\int_0^T J(X(t))\{1+|X(t)|\}dt \xrightarrow{a.s.} E[J(X(0))\{1+|X(0)|\}] < \infty \quad \text{as } T \to \infty.$$

Therefore

(4.8) $$|I_T(\theta)-I_T(\phi)| \le C^*T|\theta-\phi| \quad \text{a.s.}$$

as $T \to \infty$ for some constant $C^* > 0$. In view of (4.7) and Lemma 4.1, it follows that

(4.9) $$\frac{I_T(\theta)}{T} \xrightarrow{a.s.} I(\theta) \equiv E[f(\theta,X(0))-f(\theta_0,X(0))]^2$$

uniformly in $\theta \in \Theta$ as $T \to \infty$. But $I_T(\theta_0) = 0$ and $\lim\limits_{T\to\infty} \dfrac{I_T(\theta)}{T} > 0$ a.s. for $\theta \ne \theta_0$ by (A5). Hence, for any $\delta > 0$,

(4.10) $$\inf_{|\theta-\theta_0|\ge\delta} \frac{I_T(\theta)}{T} \xrightarrow{a.s.} \lambda \quad \text{as } T \to \infty$$

for some $\lambda > 0$ depending on δ.

Lemma 4.2. Under the assumptions (A1)-(A4), for any $T_0 > 0$ and any $\varepsilon > 0$,

(4.11) $$P(\sup_{\theta} \sup_{0\le T\le T_0} |Z_T^*(\theta)| > \varepsilon) \le C_2 \frac{T_0}{\varepsilon^2}$$

for some constant $C_2 > 0$.

Proof. Let $h(\theta,x)$ and $g(\theta,x)$ be defined as in Section 3 and

$$h(\theta,x) = \sum_n a_n(x)e^{\pi in\theta}, \quad \theta \in [-1,1].$$

Let

$$W_n^* = \int_0^T a_n(x(t))d\xi(t) .$$

Since $g(\theta,x)$ is a cubic polynomial in θ with coefficients in x, it is easy to check, by Kolmogorov's inequality, that

$$\text{(4.12)} \qquad \sup_\theta \sup_{0\le T\le T_0} \left| \int_0^T g(\theta,X(t))d\xi(t)\right| = 0_p(T_0^{\frac{1}{2}})$$

using the fact that $|\theta| \le 1$. On the other hand, for any $\varepsilon > 0$,

$$\text{(4.13)} \qquad P(\sup_\theta \sup_{0\le T\le T_0} \left|\sum_n \{\int_0^T a_n(X(t))d\xi(t)\}e^{\pi i\theta}\right| > \varepsilon)$$

$$\le P(\sup_{0\le T\le T_0} \sum_n \left|\int_0^T a_n(X(t))d\xi(t)\right| > \varepsilon)$$

$$\le \sum_n P(\sup_{0\le T\le T_0} \left|\int_0^T a_n(X(t))d\xi(t)\right| > \varepsilon_n)$$

(where $\sum\varepsilon_n \le \varepsilon$)

$$\le \sum_n \frac{1}{\varepsilon_n^2} \operatorname{Var}(\int_0^{T_0} a_n(X(t))d\xi(t))$$

(by Kolmogorov's inequality for martingales)

$$\le \sum_n \frac{1}{\varepsilon_n^2} \int_0^{T_0} E(a_n(X(t)))^2 dt$$

$$= T_0 \sum_n \frac{\mu_n}{\varepsilon_n^2}$$

$$= \frac{T_0}{\varepsilon^2} (\sum \mu_n^{1/3})^3$$

when ε_n is chosen to be $\varepsilon\mu_n^{1/3} (\sum_n \mu_n^{1/3})^{-1}$. Note that $M \equiv \sum \mu_n^{1/3} < \infty$.

Hence relations (4.12) and (4.13) together prove that

$$P(\sup_\theta \sup_{0 \le T \le T_0} |Z_T^*(\theta)| > \varepsilon) \le C_2 \frac{T_0}{\varepsilon^2} .$$

for some constant $C_2 > 0$ independent of T_0 and ε.

Lemma 4.3. For any $\gamma > 1/2$, there exists $H > 0$ such that

$$\text{(4.14)} \qquad \limsup_{T \to \infty} \sup_\theta \frac{|Z_T(\theta)|}{T^{1/2}(\log T)^\gamma} \le H \quad \text{a.s.}$$

Proof. Let

$$A_n = [\sup_{2^{n-1} < T \le 2^n} \sup_\theta |Z_T(\theta)| > H' \, 2^{n/2} n^\gamma], \quad n \ge 1.$$

Observe that Lemma 4.2 gives the inequality

$$P(A_n) = P[\sup_{0 < T \le 2^{n-1}} \sup_\theta |Z_T(\theta)| > H' 2^{n/2} n^\gamma]$$

(by stationarity of the process X(t))

$$\le \frac{C\, 2^{n-1}}{H'^2 2^n n^{2\gamma}} = \frac{C}{2H'^2} \quad \frac{1}{n^{2\gamma}} .$$

Hence $\sum_{n=1}^\infty P(A_n) < \infty$ which implies that $P(A_n$ occurs infinitely often$) = 0$ by Borel-Cantelli Lemma. Therefore $\sup_\theta |Z_T(\theta)| \le H' \, 2^{n/2} n^\gamma$ for all $2^{n-1} < T \le 2^n$ except for finitely many n with probability one and hence

$$\limsup_{T \to \infty} \sup_\theta |Z_T(\theta)| \le H \, T^{1/2} (\log T)^\gamma \quad \text{a.s.}$$

for suitable $H > 0$ depending on γ.

Theorem 4.1. Under the assumptions (A1)-(A5),

$$\hat\theta_T \to \theta_0 \quad \text{a.s. as } T \to \infty .$$

Proof. Note that

$$R_T(\theta) = I_T(\theta) + Z_T^*(\theta)$$

and $R_T(\theta_0) = 0$. Furthermore, for any $\delta > 0$, there exists $\lambda > 0$ depending on δ such that

$$\inf_{|\theta-\theta_0|\geq\delta} I_T(\theta) \geq T\lambda \quad \text{a.s.} \quad \text{as } T \to \infty$$

by (4.10) and with probability one, for any $\gamma > \frac{1}{2}$, there exists $H > 0$ depending on γ such that

$$\sup_\theta |Z_T^*(\theta)| \leq H\, T^{1/2}(\log T)^\gamma \quad \text{a.s.}$$

for sufficiently large T. Hence

$$\inf_{|\theta-\theta_0|\geq\delta} R_T(\theta) \geq \lambda^* T > 0 \quad \text{a.s.} \quad \text{as } T \to \infty.$$

for some $\lambda^* > 0$ depending on δ and γ. Since $\hat{\theta}_T$ minimizes $R_T(\theta)$ and $R_T(\theta_0) = 0$, it follows that $|\hat{\theta}_T-\theta_0| \leq \delta$ a.s. as $T \to \infty$. Hence $\hat{\theta}_T \to \theta_0$ a.s. as $T \to \infty$.

5. Asymptotic normality of the estimator

In addition to the conditions (A1)-(A5) assumed in Section 2, let us suppose that there exists a neighbourhood V_{θ_0} of θ_0 such that

(A6) $$|f_\theta^{(1)}(\theta,x)| \leq M(\theta)(1+|x|), \quad \theta \in V_{\theta_0}$$

and

$$\sup\{M(\theta): \theta \in V_{\theta_0}\} = M < \infty.$$

We shall now obtain the asymptotic distribution of $\hat{\theta}_T$ under the conditions (A1)-(A6). Since $\hat{\theta}_T$ is strongly consistent, $\hat{\theta}_T \in V_{\theta_0}$ with probability one for large T. Expanding $f(\theta,x)$ in a neighbourhood of θ_0, we have

$$f(\theta,x) = f(\theta_0,x) + (\theta-\theta_0)\tilde{f}(\theta,x)$$

where $|\tilde{\theta}-\theta_0| \le |\theta-\theta_0|$ and hence

(5.1)
$$I_T(\theta) \equiv \int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt$$
$$= (\theta-\theta_0)^2 \int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt$$
$$+ (\theta-\theta_0)^2 \int_0^T [\{f_\theta^{(1)}(\tilde{\theta},X(t))\}^2-\{f_\theta^{(1)}(\theta_0,X(t))\}^2]dt.$$

Observe that

(5.2)
$$|\{f_\theta^{(1)}(\tilde{\theta},x)\}^2-\{f_\theta^{(1)}(\theta_0,x)\}^2|$$
$$= |f_\theta^{(1)}(\tilde{\theta},x)-f_\theta^{(1)}(\theta_0,x)||f_\theta^{(1)}(\tilde{\theta},x)+f_\theta^{(1)}(\theta_0,x)|$$
$$\le 2\, M|\tilde{\theta}-\theta_0|^\alpha\, c(x)(1+|x|)$$

by assumptions (A3) and (A6). Therefore

(5.3)
$$|I_T(\theta) - (\theta-\theta_0)^2 \int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt|$$
$$\le 2\, M|\theta-\theta_0|^{2+\alpha} \int_0^T c(X(t))(1+|X(t)|)dt.$$

Let us write $\theta-\theta_0 = T^{-1/2}\psi$. Then it follows that

(5.4)
$$\sup_{|\psi|\le A_T} |I_T(\theta)-\psi^2T^{-1}\int_0^T\{f_\theta^{(1)}(\theta_0,X(t))\}^2dt| \le M_1A_T^{2+\alpha}T^{-1-\alpha}$$

for some constant $M_1 > 0$ by the erogodic theorem since

$$E(c(x(0))(1+|X(0)|) < \infty.$$

On the other hand, let

$$v_T(\psi,x) = T^{1/2}[f(\theta_0+\psi T^{-1/2},x)-f(\theta_0,x)-\psi T^{-1/2}f_\theta^{(1)}(\theta_0,x)]$$

for $|\psi| \leq A_T$. Then $v_T(\psi,x)$ is differentiable with respect to ψ and the derivative $v_T^{(1)}(\psi,x)$ satisfies

$$v_T^{(1)}(\psi,x)-v_T^{(1)}(\zeta,x) = f_\theta^{(1)}(\theta_0+\psi T^{-1/2},x)-f_\theta^{(1)}(\theta_0+\zeta T^{-1/2},x)$$

and hence

$$(5.5) \qquad |v_T^{(1)}(\psi,x)-v_T^{(1)}(\zeta,x)| \leq c(x)T^{-\alpha/2}|\psi-\zeta|^\alpha$$

by (A3) for all ψ, ζ in $[-A_T,A_T]$. It can be shown that there exists a polynomial in ψ with coefficients in x viz

$$(5.6) \qquad g_T(\psi,x) = v_T(A_T,x)P_1(\tfrac{\psi}{A_T}) + A_T v_T^{(1)}(A_T,x)P_2(\tfrac{\psi}{A_T})$$

$$+ v_T(-A_T,x)P_3(\tfrac{\psi}{A_T}) + A_T v_T^{(1)}(-A_T,x)P_4(\tfrac{\psi}{A_T})$$

on $[-A_T,A_T]$ such that

$$(5.7) \qquad g_T(A_T,x) = v_T(A_T,x), g_T(-A_T,x) = v_T(-A_T,x),$$

$$(5.8) \quad g_T^{(1)}(A_T,x) = v_T^{(1)}(A_T,x) \text{ and } g_T^{(1)}(-A_T,x) = v_T^{(1)}(-A_T,x)$$

where P_i, $1 \leq i \leq 4$ are polynomials in $\frac{\psi}{A_T}$ with constant coeffiecients. Observing that $v_T(0,x) = v_T^{(1)}(0,x) = 0$, it is easy to check that

$$(5.9) \qquad |g_T^{(1)}(A_T,x)| \leq c(x)A_T^\alpha T^{-\alpha/2} \quad ,$$

$$(5.10) \qquad |g_T^{(1)}(-A_T,x)| \leq c(x)A_T^\alpha T^{-\alpha/2} \quad ,$$

$$(5.11) \qquad |g_T(A_T,x)| \leq c(x)A_T^{1+\alpha}T^{-\alpha/2} \quad ,$$

and

$$(5.12) \qquad |g_T(-A_T,x)| \leq c(x)A_T^{1+\alpha}T^{-\alpha/2} \quad .$$

Furthermore there exists a constant $M_2 > 0$ independent of T such that

$$(5.13) \qquad |g_T^{(1)}(\psi,x)-g_T^{(1)}(\zeta,x)| \leq M_2 c(x)A_T^{\alpha-1}T^{-\alpha/2}|\psi-\zeta|$$

for all $\psi,\zeta \in [-A_T,A_T]$. But

$$A_T^{\alpha-1}|\psi-\zeta| \leq 2^{1-\alpha}|\psi-\zeta|^\alpha$$

since $|\psi-\zeta| \leq 2A_T$. Hence there exists a constant $M_3 > 0$ independent of T such that

$$|g_T^{(1)}(\psi,x)-g_T^{(1)}(\zeta,x)| \leq M_3c(x)T^{-\alpha/2}|\psi-\zeta|^\alpha \tag{5.14}$$

for all ψ, $\zeta\in[-A_T,A_T]$. Renormalizing, we get that

$$|g_T^{(1)}(\psi^*,x)-g_T^{(1)}(\zeta^*,x)| \leq M_3c(x)A_T^\alpha|\psi^*-\zeta^*|^\alpha T^{-\alpha/2} \tag{5.15}$$

for all $\psi^*,\zeta^* \in [-1,1]$. Let

$$h_T(\psi^*,x) = v_T(\psi^*,x)-g_T(\psi^*,x). \tag{5.16}$$

Then there exists a constant $M_3^* > 0$ independent of T such that

$$|h_T^{(1)}(\psi^*,x)-h_T^{(1)}(\zeta^*,x)| \leq M_3^*c(x)A_T^\alpha|\psi^*-\zeta^*|^\alpha T^{-\alpha/2} \tag{5.17}$$

for all $\psi^*,\zeta^* \in [-1,1]$ by relations (5.5) and (5.15). Now, applying Fourier series methods as in Lemma 4.2, it can be shown that for every $\varepsilon > 0$,

$$P(\sup_{|\psi^*|\leq 1} |\int_0^T v_T(\psi^*,X(t))d\xi(t)| > \varepsilon) \leq \frac{M_4T}{\varepsilon^2} A_T^{2\alpha}T^{-\alpha}E[c^2(X(0)]$$

and hence

$$P(\sup_{|\psi|\leq A_T} |\int_0^T \{f(\theta_0+\psi T^{-1/2},X(t))-f(\theta_0,X(t)) - \psi T^{-1/2}f_\theta^{(1)}(\theta_0,X(t))\}d\xi(t)| > \varepsilon) \leq \frac{M_4}{\varepsilon^2} A_T^{2\alpha}T^{-\alpha}E[c^2(X(0)] . \tag{5.18}$$

Let us choose $A_T = \log T$. Since

$$\frac{1}{T}\int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt \to I(\theta_0) = E[f_\theta^{(1)}(\theta_0,X(0))]^2 \quad \text{a.s.}$$

as $T \to \infty$ by the ergodic theorem and

$$\frac{1}{\sqrt{T}}\int_0^T f_\theta^{(1)}(\theta_0,X(t))d\xi(t) \xrightarrow{\mathcal{L}} N(0,\sigma^2 I(\theta_0)) \quad \text{as } T \to \infty$$

by the central limit theorem for stochastic integrals (cf. Basawa and Prakasa Rao (1979)), relations(5.4) and (5.18) imply that the asymptotic distribution of $\hat{\theta}_T$ which minimizes $R_T(\theta)$ given by (2.9) can be obtained from the process

(5.19) $$\psi^2 I(\theta_0) - 2\psi Z, \quad -\infty < \psi < \infty$$

where Z is normal with mean 0 and variance $\sigma^2 I(\theta_0)$. Since

$$\hat{\psi} = Z/I(\theta_0)$$

minimizes (5.16), it follows that

(5.20) $$T^{1/2}(\hat{\theta}_T-\theta_0) \xrightarrow{\mathcal{L}} N(0,\sigma^2/I(\theta_0)).$$

This result is obtained under stronger conditions in Prakasa Rao (1979b) for the least squares estimator $\hat{\theta}_{n,T}$ defined at the beginning of Section 2.

Appendix

Lemma 1 Suppose $\phi(u)$ is square integrable on [-1,1] and $\phi(\cdot)$ is Lipschitz of order α i.e., then exists c > 0 such that

(1) $$|\phi(u)-\phi(v)| \leq c|u-v|^\alpha .$$

Let $\phi(u) = \sum_n a_n e^{\pi i n u}$. Then for any $0 < \gamma < \alpha$,

(2) $$\sum_n |a_n|^2 n^{2\gamma} \leq K_1(\alpha,\gamma)c^2 .$$

Proof. It is easy to check that

(3) $$\int_{-1}^{1} |\phi(u+h)-\phi(u-h)|^2 du = 4 \sum_n |a_n|^2 \sin^2 \pi nh.$$

Since ϕ is Lipschitz satisfying (1), it follows that

(4) $$4 \sum_n |a_n|^2 \sin^2 \pi nh \le 2^{2\alpha+1} c^2 h^{2\alpha}$$

for all $h \in [0,1]$. Let $h = 2^{-k}$ and $2^{k-2} < n \le 2^{k-1}$. It is clear that $\sin^2 \pi nh \ge \frac{1}{2}$ and relation (4) shows that

(5) $$\sum_{n=2^{k-2}+1}^{2^{k-1}} |a_n|^2 \le 2^{2\alpha} c^2 2^{-2k\alpha}$$

for any $k \ge 2$ and hence for any $0 < \gamma < \alpha$,

(6) $$\sum_{n=2^{k-2}+1}^{2^{k-1}} |a_n|^2 n^{2\gamma} \le 2^{2\alpha} c^2 2^{(2\gamma-2\alpha)k} .$$

Summing over all $k \ge 2$, we obtain that

(7) $$\sum_n |a_n|^2 n^{2\gamma} \le 2^{2\alpha} c^2 (1-2^{(2\gamma-2\alpha)})^{-1} .$$

Hence there exists a constant $K_1(\alpha,\gamma) > 0$ such that

(8) $$\sum_n |a_n|^2 n^{2\gamma} \le K_1(\alpha,\gamma) c^2$$

where c is the Lipschitzian constant given by (1).

Remark. A slight variation of the above result is due to Szasz (1922). The proof given above is the same as in Szasz (1922) and is given here for completeness.

Lemma 2. Suppose $h(u)$ is square integrable on $[-1,1]$ with $h(-1)=h(1)=0$ and $h'(\cdot)$ exists and is Lipschitzian of order α i.e., there exists $c > 0$ such that

(9) $$|h'(u)-h'(v)| \le c|u-v|^{\alpha}.$$

Let $h(u) = \sum_n a_n e^{\pi i n u}$. Then, for any $0 < \gamma < \alpha$,

(10) $$\sum_n |a_n|^2 n^{2+2\gamma} \le K_2(\alpha,\gamma)c^2$$

and

(11) $$\sum_n |a_n|^{2/3} \le K_3(\alpha,\gamma)c^2 .$$

Proof. Since $h'(u) = \pi i \sum_n n a_n e^{\pi i n u}$, inequality (10) follows from Lemma 1. Observe that

$$\sum_n |a_n|^{2/3} \le (\sum |a_n|^2 n^{2+2\gamma})^{1/3} (\sum n^{-(1+\gamma)})^{2/3}$$

$$\le K_2(\alpha,\gamma)c^2 (\sum n^{-(1+\gamma)})^{2/3}$$

$$= K_3(\alpha,\gamma)c^2 .$$

Lemma 3. Let $h(\theta,x) = \sum_n a_n(x)e^{\pi i n\theta}$ and suppose there exists $\alpha > 0$ such that

$$|h_\theta^{(1)}(\theta,x)-h_\theta^{(1)}(\phi,x)| \le c(x)|\theta-\phi|^{\alpha}$$

for all θ, ϕ in $[-1,1]$ where $f_\theta^{(1)}$ denotes the partial derivative of f with respect to θ. Let $\{X(t),\ t \in [0,T]\}$ be a stochastic process such that

$$E[h(\theta,X(t))]^2 < \infty$$

for every $t \in [0,T]$. Then, for any $\gamma < \alpha$, there exists a positive constant $K_4(\alpha,\gamma)$ such that

$$\sum_n \{\frac{1}{T}\int_0^T E[a_n^2(X(t))]dt\}^{1/3} \le K_4(\alpha,\gamma)\{\frac{1}{T}\int_0^T E(c^2(X(t))dt\}^{1/3}.$$

Proof. By Lemma 2, it follows that

$$\sum_n |a_n(X(t))|^2 n^{2+2\gamma} \leq K_2(\alpha,\gamma)c^2(X(t)) \quad \text{a.s.}$$

for every $t \in [0,T]$. Hence

$$\sum_n E[a_n^2(X(t))]n^{2+2\gamma} \leq K_2(\alpha,\gamma)E[c^2(X(t))]$$

for all $t \in [0,T]$. Let

$$\mu_n = \frac{1}{T}\int_0^T E[a_n^2(X(t))]dt.$$

The inequality proved above gives the relation

$$\sum_n \mu_n n^{2+2\gamma} \leq K_2(\alpha,\gamma)\frac{1}{T}\int_0^T E[c^2(X(t))]dt$$

and hence

$$\sum_n \mu_n^{1/3} \leq (\sum \mu_n n^{2+2\gamma})^{1/3}(\sum n^{-(1+\gamma)})^{2/3}$$

$$\leq K_2^{1/3}(\alpha,\gamma)(\sum n^{-(1+\gamma)})^{2/3}\{\frac{1}{T}\int_0^T E[c^2(X(t))]dt\}^{1/3}$$

$$\leq K_4(\alpha,\gamma)\{\frac{1}{T}\int_0^T E[c^2(X(t))]dt\}^{1/3}.$$

Remark. Analgous argument proves that

$$\sum \mu_n^{1/2} \leq (\sum \mu_n \cdot n^{2+2\gamma})^{1/2}(\sum n^{-2(1+\gamma)})^{1/2}$$

$$< \infty .$$

Acknowledgement. One of the authors (B.L.S.Prakasa Rao) thanks the Departments of Statistics and Mathematics of Purdue University for inviting him to spend the Summer 1979 which made the colloboration possible.

REFERENCES

[1] Basawa, I.V., Prakasa Rao, B.L.S. (1979). Statistical Inference for Stochastic Processes, Theory and Methods, Academic Press, London. (To appear).

[2] Dorogovcev, A. Ja. (1976). The consistency of an estimate of a parameter of a stochastic differential equaiton, Theory of Probability and Math. Statist. 10, 73-82.

[3] Gikhman, I.I. and Skorokhod, A.V. (1972). Stochastic Differential Equations, Springer-Verlag, Berlin.

[4] McKean, H. P. (1969). Stochastic Integrals, Academic Press, New York.

[5] Prakasa Rao, B.L.S. (1979a). The Bernstein-von Mises theorem for a class of diffusion processes, (Preprint) The University of Poona.

[6] Prakasa Rao, B.L.S. (1979b). Asymptotic theory for non-linear least squares estimators for diffusion processes, (Preprint), Indian Statistical Institute, New Delhi.

[7] Szász, O. (1922). Über den Kongvergenzexponent der Fourierschen Reihen, Münchener Sitzungskerichte, 135-150.

Department of Statistics
Purdue University
West Lafayette, IN 47907

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Mimeograph Series #79-13	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Asymptotic Theory for Process Least Squares Estimators for Diffusion Processes		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Mime. Series #79-13
7. AUTHOR(s) B.L.S.Prakasa Rao and Herman Rubin		8. CONTRACT OR GRANT NUMBER(s) ONR-N00014-75-C-0455 NSF-MCS76-08316
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University West Lafayette, Indiana 47907 99173J		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, D.C.		12. REPORT DATE July 1979
		13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stochastic Differential Equation; Diffusion

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Strong consistency and asymptotic normality of an estimator related to least squares estimator for parameters involved in nonlinear stochastic differential equations are investigated by studying families of stochastic integrals using Fourier analytic methods.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A073 577 PURDUE UNIV LAFAYETTE IND DEPT OF STATISTICS F/G 12/1

ASYMPTOTIC THEORY FOR PROCESS LEAST SQUARES ESTIMATORS FOR DIFF--ETC(U)

JUL 79 B L RAO, H RUBIN N00014-75-C-0455

UNCLASSIFIED MMS-79-13 NL

1 OF 1

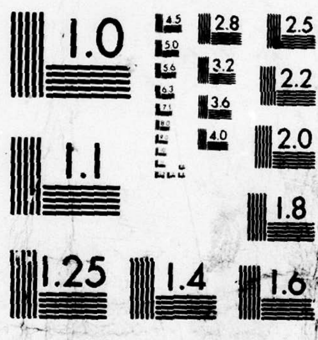
AD A073577

END

DATE FILMED

10-79

DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1

(12) LEVEL II

AD A073577

PURDUE UNIVERSITY

DDC
RECEIVED
SEP 7 1979
B

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

79 09 5 008

(12) LEVEL II

(6)

ASYMPTOTIC THEORY FOR PROCESS LEAST SQUARES ESTIMATORS FOR DIFFUSION PROCESSES.

by

(10) B.L.S. Prakasa Rao* and Herman Rubin**
Indian Statistical Institute, New Delhi
and
Purdue University

Department of Statistics
Division of Mathematical Sciences
(9) Mimeograph Series # 79-13 (14) MMS-

(11) July 1979

(12) 29p.

(15) * Research partially supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. NSF-MCS76-08316

** Research partially supported by NSF Grant MCS76-08316.

291 730

ABSTRACT

Strong consistency and asymptotic normality of an estimator related to least squares estimator for parameters involved in nonlinear stochastic differential equations are investigated by studying families of stochastic integrals using Fourier analytic methods.

AMS (1980) Subject Classification: Primary 62M05, Secondary 60H10

Key words and phrases: Stochastic Differential Equation; Diffusion Process; Least Squares Estimation; Consistency; Asymptotic Normality

1. Introduction

Recently there is a growing interest in the study of inference problems for stochastic processes both continuous and discrete time in view of the large number of applications to engineering problems. It has been found that the class of diffusion processes is amenable for statistical analysis. A survey of the recent work in this area is given in Basawa and Prakasa Rao (1979). Further work on asymptotic theory of maximum likelihood and Bayes estimators for parameters of diffusion processes is discussed in Prakasa Rao (1979a).

Dorogovchev (1976) studied weak consistency of least square estimators for parameters of diffusion processes which are solutions of non-linear stochastic differential equations. Asymptotic normality and asymptotic effiency of these estimators is investigated in Prakasa Rao (1979b). Our aim in this paper is to study limiting properties of a process related to least squares estimator and hence to discuss the asymptotic properties of an estimator derived from the limiting process. We study strong consistency and asymptotic normality of this estimator. Our approach here is entirely different from that of Dorogovchev (1976) and Prakasa Rao (1979b). We believe that our techniques for study of families of stochastic integrals is new and is of independent interest.

2. Study of process related to least squares estimator

Let $\{X(t), t \geq 0\}$ be a real-valued stationary ergodic process satisfying the stochastic differential equation

$$dX(t) = f(\theta_0, X(t))dt + d\xi(t), \quad X(0) = X_0, \quad t \geq 0$$

where $\xi(t)$ is a Wiener process with mean zero and variance $\sigma^2 t$, σ^2 being unknown and $E[X_0^2] < \infty$. Suppose $f(\theta, x)$ is a known real-valued function continuous on $\Theta \times R$ where Θ is a closed interval on the real line and $\theta_0 \in \Theta$ is unknown. Without loss of generality, assume that $\Theta = [-1,1]$.

Suppose the process $\{X(t), 0 \leq t \leq T\}$ is observed at time points t_k, $k = 0,1,\ldots,n-1$ with $t_0 = 0$ and $t_n = T$. Let

$$Q_n^T(\theta) = \sum_{k=0}^{n-1} \frac{[X(t_{k+1}) - X(t_k) - f(\theta, X(t_k))\Delta t_k]^2}{t_k}.$$

where $\Delta t_k = t_{k+1} - t_k$, $\theta \leq k \leq n-1$. An estimator $\hat{\theta}_{n,T}$ which minimizes $Q_n^T(\theta)$ over $\theta \in \Theta$ is called a <u>least squares estimator</u> of θ. Assume that such an estimator exists. Note that if $\hat{\theta}_{n,T}$ minimizes $Q_n^T(\theta)$, then it minimizes $Q_n^T(\theta) - Q_n^T(\theta_0)$.

We shall first study the limiting properties of the process $\{Q_n^T(\theta) - Q_n^T(\theta_0), \ \theta \in \Theta\}$ as the norm of division $\Delta_n = \max_{1 \leq k \leq n} |t_{k+1} - t_k|$ tends to zero. Let $\Delta X_k = X(t_{k+1}) - X(t_k)$ and $\Delta\xi_k = \xi(t_{k+1}) - \xi(t_k)$, $0 \leq k \leq n-1$. For any fixed θ,

$$Q_n^T(\theta)-Q_n^T(\theta_0)$$

$$= \sum_k \frac{1}{\Delta t_k}\,[\Delta X_k - f(\theta, X(t_k))\Delta t_k]^2$$

$$- \sum_k \frac{1}{\Delta t_k}\,[\Delta X_k - f(\theta_0, X(t_k))\Delta t_k]^2$$

$$= \sum_k \frac{1}{\Delta t_k}\,\{\int_{t_{k-1}}^{t_{k+1}} f(\theta_0, X(t))dt + \Delta\xi_k - f(\theta, X(t_k))\Delta t_k\}^2$$

$$- \sum_k \frac{1}{\Delta t_k}\,\{\int_{t_k}^{t_{k+1}} f(\theta_0, X(t))dt + \Delta\xi_k - f(\theta_0, X(t_k))\Delta t_k\}^2$$

$$= \sum_k \frac{1}{\Delta t_k}\,\{\int_{t_k}^{t_{k+1}} [f(\theta_0, X(t)) - f(\theta, X(t_k))]dt + \Delta\xi_k\}^2$$

$$- \sum_k \frac{1}{\Delta t_k}\,\{\int_{t_k}^{t_{k+1}} [f(\theta_0, X(t)) - f(\theta_0, X(t_k))]dt + \Delta\xi_k\}^2 .$$

It is easy to check that

$$(2.0)\quad Q_n^T(\theta)-Q_n^T(\theta_0)$$

$$= \sum_k [f(\theta_0, X(t_k)) - f(\theta, X(t_k))]^2 \Delta t_k$$

$$+ 2\sum_k [f(\theta_0, X(t_k) - f(\theta, X(t_k)]\Delta\xi_k$$

$$+ 2\sum_k \{f(\theta_0, X(t_k) - f(\theta, X(t_k))\} \int_{t_k}^{t_{k+1}} \{f(\theta_0, X(t)) - f(\theta_0, X(t_k))\}dt$$

$$= I_{1n} + 2I_{2n} + 2I_{3n} .$$

Assume that the regularity condition on $f(x,\theta)$ stated at the end of this section are satisfied. Since $f(\theta,x)$ is continuous in x and the

process X has continuous sample paths with probability one, it follows that

$$(2.1) \qquad I_{1n} \xrightarrow{a.s.} \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]^2 dt$$

as $\Delta_n \to 0$. Assumption (A2) implies that

$$(2.2) \qquad I_{2n} \xrightarrow{q.m.} \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]d\xi(t)$$

as $\Delta_n \to 0$ in view of stationarity of the process X where the last integral is the Ito-stochastic integral.

Let us now estimate I_{3n}. In view of assumption (A4), it can be checked that

$$(2.3) \qquad \left|\int_{t_k}^{t_{k+1}} \{f(\theta_0,X(t))-f(\theta_0,X(t_k))\}dt\right|$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} |X(t)-X(t_k)|dt$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} \{|\xi(t)-\xi(t_k)|+\int_{t_k}^{t} f(\theta_0,X(s))|ds\}dt$$

$$\leq L(\theta_0)\int_{t_k}^{t_{k+1}} |\xi(t)-\xi(t_k)|dt+L^2(\theta_0)\int_{t_k}^{t_{k+1}} \left[\int_{t_k}^{t}\{1+|X(s)|\}ds\right] dt$$

$$\leq L(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} |\xi(t)-\xi(t_k)|+L^2(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} \int_{t_k}^{t} \{1+|X(s)|\}ds.$$

$$\leq L(\theta_0)\Delta t_k \sup_{t_k\leq t\leq t_{k+1}} |\xi(t)-\xi(t_k)|+L^2(\theta_0)\Delta t_k^2 \sup_{t_k\leq t\leq t_{k+1}} \{1+|X(t)|\}$$

for $0 \leq k \leq n-1$. Using assumption (A4) again, we obtain the following inequality:

$$(2.4)\qquad I_{3n} \le \sum_k J(X(t_k))\Big\{L(\theta_0)\cdot\Delta t_k \sup_{t_k\le t\le t_{k+1}} |\xi(t)-\xi(t_k)| + L^2(\theta_0)\Delta t_k^2 \sup_{t_k\le t\le t_{k+1}} \{1+|X(t)|\}\Big\}|\theta-\theta_0|.$$

Since $J(\cdot)$ is continuous and $X(\cdot)$ has continuous sample paths almost surely, it follows that there exists a constant $C^*(\theta_0)$ depending on T only such that

$$(2.5)\qquad I_{3n} \le C^*(\theta_0)\Big\{\sum_k \Delta t_k\cdot \sup_{t_k\le t\le t_{k+1}} |\xi(t)-\xi(t_k)| + \sum_k \Delta t_k^2\Big\}|\theta-\theta_0| .$$

Since $\theta\in\Theta$ compact, it follows that

$$I_{3n} \le C(\theta_0)\{\sum_k \Delta t_k(1+\Delta t_k)(2\Delta t_k \log 1/\Delta t_k)^{1/2} + \sum_k \Delta t_k^2\} \quad \text{a.s.}$$

whenever Δ_n is sufficiently small by the law of iterated logarithm for Brownian increments (cf. McKean (1969), p.14). Therefore

$$(2.6)\qquad I_{3n} = O(\sum_k \Delta t_k^{3/2} \log^{1/2} 1/\Delta t_k) \text{ a.s.}$$

uniformly in $\theta\in\Theta$. Furthermore the convergence in (2.1) is uniform in $\theta\in\Theta$ since

$$|f(\theta_0,X(t))-f(\theta,X(t))|^2 \le |\theta_0-\theta|^2 J^2(X(t)) \le C\, J^2(X(t))$$

and $J(X(t))$ is integrable pathwise on $[0,T]$ by (A4). Here we have used the fact that Θ is compact. Hence

$$(2.7)\qquad I_{1n} = \int_0^T [f(\theta_0,X(t)-f(\theta,X(t)]^2 dt + o(1) \qquad \text{a.s.}$$

uniformly in θ as $\Delta_n \to 0$. We shall discuss uniform convergence of I_{2n} in the next section.

Relations (2.0), (2.6) and (2.7) show that, for any fixed T,

$$(2.8) \quad Q_n^T(\theta)-Q_n^T(\theta_0) = \int_0^T [f(\theta_0,X(t))-f(\theta,X(t))]^2 dt + I_{2n}+o(1) \text{ a.s.}$$

uniformly in $\theta \in \Theta$ compact as $\Delta_n \to 0$ where I_{2n} satisfies relation (2.2). Let us consider the limiting process

$$(2.9) \quad R_T(\theta) = \int_0^T [f(\theta_0,X(t)) - f(\theta,X(t))]^2 dt$$

$$+ 2\int_0^T [f(\theta_0,X(t)) - f(\theta,X(t))]d\xi(t)$$

$$= \int_0^T v^2(\theta,X(t))dt - 2\int_0^T v(\theta,X(t))d\xi(t)$$

where

$$(2.10) \quad v(\theta,x) = f(\theta,x) - f(\theta_0,x).$$

We study the limiting properties of the process $\{R_T(\theta), \theta \in \Theta\}$ in the next section.

Assumptions

(A1) $f(\theta,x)$ is continuous in (θ,x) and differentiable with respect to θ. Denote the first partial derivative of f with respect to θ by $f_\theta^{(1)}(\theta.x)$ and the derivative evaluated at θ_0 by $f_\theta^{(1)}(\theta_0,x)$.

(A2) $E[f_\theta^{(1)}(\theta_0,X(0))]^2 < \infty$

(A3) $f_\theta^{(1)}(\theta,x)$ is Lipschitzian in θ for each x i.e., there exists $\alpha > 0$ such that

$$|f_\theta^{(1)}(\theta,x)-f_\theta^{(1)}(\phi,x)| \leq c(x)|\theta-\phi|^\alpha, \quad x \in R,\theta,\phi \in \Theta$$

and

$$E[c^2(X(0))] < \infty.$$

(A4) $f(\theta,x)$ satisfies the following conditions:

(i) $|f(\theta,x)| \le L(\theta)(1+|x|)$, $\theta \in \Theta$, $x \in R$; $\sup\{L(\theta): \theta \in \Theta\} < \infty$.

(ii) $|f(\theta,x)-f(\theta,y)| \le L(\theta)|x-y|$, $\theta \in \Theta$, $x,y, \in R$.

(iii) $|f(\theta,x)-f(\phi,x)| \le J(x)|\theta-\phi|$, $\theta , \phi \in \Theta$, $x \in R$

where $J(\cdot)$ is continuous and $E[J^2(X(0))] < \infty$.

(A5) $I(\theta) \equiv E[f(\theta,X(0))-f(\theta_0,X(0)]^2 > 0$ for $\theta \neq \theta_0$.

Remark: Since $E[X^2(0)] < \infty$, assumption A4(i) implies that

$$E[f(\theta,X(0))]^2 < \infty$$

for all $\theta \in \Theta$.

3. Study of a limiting process related to least squares estimator

Let us now study the properties of the limiting process

$$Z_T(\theta) \equiv \frac{1}{\sqrt{T}} \int_0^T v(\theta,X(t))d\xi(t) \tag{3.1}$$

as a process in the parameter $\theta \in \Theta = [-1,1]$ as $T \to \infty$. From the central limit theorem for stochastic integrals (cf. Basawa and Prakasa Rao (1979)), it can be shown that

$$\frac{1}{\sqrt{T}} \int_0^T v(\theta,X(t))d\xi(t) \xrightarrow{\mathcal{L}} N(0,E[v(\theta,X(0))]^2\sigma^2)$$

since the process X is stationary ergodic. In general, finite dimensional distributions of the process $\{Z_T(\theta), \theta \in \Theta\}$ converge to the finite dimensional distributions of the Gaussian process $\{Z(\theta), \theta \in \Theta\}$ with mean zero and covariance function

$$R(\theta_1,\theta_2) = E[v(\theta_1,X(0))v(\theta_2,X(0))]\sigma^2.$$

We shall now prove the weak convergence of the process $\{Z_T(\theta), \theta \in \Theta\}$ on C[-1,1] under uniform norm. It is sufficient to prove that

$$\lim_{T\to\infty} \overline{\lim_{\delta\to 0}} P(\sup_{|\theta-\phi|<\delta} |Z_T(\theta)-Z_T(\phi)| > \varepsilon) = 0. \tag{3.2}$$

Since $v(\theta,x)$ is differentiable with respect to θ on $[-1,1]$ by assumption (A1), it is easy to see that there exists a cubic polynomial $g(\theta,x)$ in θ such that

$$g(-1,x) = v(-1,x),\ g(1,x) = v(1,x)$$

and

$$g_\theta^{(1)}(-1,x) = v_\theta^{(1)}(-1,x),\ g_\theta^{(1)}(1,x) = v_\theta^{(1)}(1,x).$$

Let

$$h(\theta,x) = v(\theta,x)-g(\theta,x).$$

Then $h(-1,x) = h(1,x) = 0, h_\theta^{(1)}(-1,x) = h_\theta^{(1)}(1,x) = 0$. Now

$$Z_T(\theta) = \frac{1}{\sqrt{T}} \int_0^T h(\theta,x(t))d\xi(t) + \frac{1}{\sqrt{T}} \int_0^T g(\theta,X(t))d\xi(t). \tag{3.3}$$

Since $g(\theta,x)$ is a cubic polynomial in θ with coefficients in x which are linear functions of $v(-1,x)$, $v(1,x)$, $v_\theta^{(1)}(-1,x)$ and $v_\theta^{(1)}(1,x)$, it is easy to check the uniform equi-continuity condition of type (3.2) for

$$\frac{1}{\sqrt{T}} \int_0^T g(\theta,X(t))d\xi(t).$$

Let us now consider the process

$$W_T(\theta) = \frac{1}{\sqrt{T}} \int_0^T h(\theta,X(t))d\xi(t). \tag{3.4}$$

Let the Fourier expansion for $h(\theta,x)$ in $L_2([-1,1])$ be given by

$$h(\theta,x) = \sum_n a_n(x)e^{\pi i n\theta}, \qquad x \in R. \tag{3.5}$$

<u>Lemma 3.1</u>

$$\int_0^T h(\theta,X(t))d\xi(t) = \sum_n \{ \int_0^T a_n(X(t))d\xi(t)\}e^{\pi i n\theta} \tag{3.6}$$

in the sense of convergence in quadratic mean.

Proof An approximating sum in L_2-norm for

$$\int_0^T h(\theta,X(t))d\xi(t)$$

is

$$A_{1N} = \sum_{j=1}^{N} h(\theta,X(t_{j-1}))\Delta\xi_j$$

and an approximating sum in L_2-norm for $\sum_n \{\int_0^T a_n(X(t))d\xi(t)\}e^{\pi in\theta}$ is

$$A_{2NM} = \sum_{|n|\leq M} e^{\pi in\theta}(\sum_{j=1}^{N} a_n(X(t_{j-1})\Delta\xi_j) .$$

It is sufficient to prove that $E|A_{1N}-A_{2NM}|^2 \to 0$ as $N \to \infty$ and $M \to \infty$. Now

$$E|A_{1N}-A_{2NM}|^2 = E|\sum_{j=1}^{N}\{(h(\theta,X(t_{j-1}))-\sum_{n=-M}^{M} e^{\pi in\theta}a_n(X(t_{j-1}))\}\Delta\xi_j|^2$$

$$= E|\sum_{j=1}^{N}\sum_{|n|>M} a_n(X(t_{j-1}))e^{\pi in\theta}\Delta\xi_j|^2$$

$$\leq [\sum_{|n|>M}\{E(\sum_{j=1}^{N} a_n(X(t_{j-1}))\Delta\xi_j)^2\}^{\frac{1}{2}}]^2$$

by the elementary inequality

$$E|\sum_n \lambda_n Y_n|^2 \leq (\sum_n|\lambda_n|(E(Y_n^2))^{\frac{1}{2}})^2$$

for any sequence of complex numbers $\{\lambda_n\}$ and any sequence of real valued random variables $\{Y_n, n \geq 1\}$. Hence

$$E|A_{1N}-A_{2NM}|^2 \leq [\sum_{|n|>M}\{\sum_{j=1}^{N} E(a_n(X(t_{j-1}))^2\Delta t_j\}^{\frac{1}{2}}]^2 .$$

Since

$$\sum_{j=1}^{N} E(a_n(X(t_{j-1}))^2\Delta t_j \to \int_0^T E\{a_n(X(t)\}^2dt = T\mu_n \quad \text{(say)},$$

as $N \to \infty$, it is sufficient to prove that $\sum_n \mu_n^{\frac{1}{2}} < \infty$. This follows from remarks following Lemma 3 of the appendix under assumption (A3).

Let

$$W_n = \frac{1}{\sqrt{T}} \int_0^T a_n(X(t))d\xi(t). \tag{3.7}$$

Lemma 3.2. For every $\varepsilon > 0$,

$$\lim_{\delta\to 0} P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big) = 0 \tag{3.8}$$

for every $T > 0$.

Proof. In view of Lemma 3.1, for any $\varepsilon > 0$,

$$\begin{aligned} (3.9)\quad & P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big) \\ &= P\Big(\sup_{|\theta-\phi|<\delta} \Big| \sum_{n=-\infty}^{\infty} W_n(e^{\pi in\theta}-e^{\pi in\phi})\Big| > \varepsilon\Big) \\ &\le P\Big(\sup_{|\theta-\phi|<\delta} \sum_{n=-\infty}^{\infty} |W_n|\, |e^{\pi in\theta}-e^{\pi in\phi}| > \varepsilon\Big). \end{aligned}$$

Let n_0 be chosen so that

$$\sum_{n=n_0}^{\infty} \mu_n^{1/3} < \varepsilon 2^{-4/3}. \tag{3.10}$$

This is possible since $\sum_{n=1} \mu_n^{1/3} < \infty$ by Lemma 3 of the appendix.

Inequality (3.9) implies that

$$P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big)$$

$$\leq P\Big(\sup_{|\theta-\phi|<\delta} \sum_{n=-n_0}^{n_0} |W_n|\,|n|\,|\theta-\phi| > \frac{\varepsilon}{4\pi}\Big) + P\Big(\sum_{|n|>n_0} |W_n| > \frac{\varepsilon}{2}\Big)$$

$$\leq \sum_{n=1}^{n_0} P\Big(|W_n| > \frac{\varepsilon}{2\pi n_0\delta}\Big) + 2\sum_{n=n_0+1}^{\infty} P(|W_n| > \varepsilon_n)$$

$$\Big(\text{Here } \varepsilon_n = \frac{\varepsilon}{2^{4/3}}\,\mu_n^{1/3}\Big(\sum_{n=n_0+1}^{\infty} \mu_n^{1/3}\Big)^{-1}\Big)$$

$$\leq \Big(\frac{2\pi n_0\delta}{\varepsilon}\Big)^2 \sum_{n=1}^{n_0} \mu_n + \sum_{n=n_0+1}^{\infty} \frac{\mu_n}{\varepsilon_n^2}$$

$$(\text{since } E(W_n) = 0 \text{ and } \operatorname{Var}(W_n) = \mu_n)$$

$$= \frac{(2\pi n_0\delta)^2}{\varepsilon^2} \sum_{n=1}^{n_0} \mu_n + \frac{8}{\varepsilon^2}\Big(\sum_{n=n_0+1}^{\infty} \mu_n^{1/3}\Big)^3$$

$$= C_{n_0}\frac{\delta^2}{\varepsilon^2} + \frac{8}{\varepsilon^2}\Big(\frac{\varepsilon}{2}\Big)^3$$

where C_{n_0} depends only on n_0. Choosing δ such that

$$C_{n_0}\frac{\delta^2}{\varepsilon^2} < \varepsilon \quad \text{i.e.} \quad 0 < \delta < \left(\frac{\varepsilon^3}{2C_{n_0}}\right)^{\frac{1}{2}}$$

we have the inequality

$$P\Big(\sup_{|\theta-\phi|<\delta} |W_T(\theta)-W_T(\phi)| > \varepsilon\Big) \leq 2\varepsilon$$

for every $0 < \delta < \left(\frac{\varepsilon^3}{2C_{n_0}}\right)^{\frac{1}{2}}$ and for every $T > 0$. This proves (3.8).

Theorem 3.1. The family of stochastic processes $\{Z_T(\theta), \theta \in \Theta\}$ on $C[-1,1]$ converge in distribution to the Gaussian process with mean zero and covariance function

$$R(\theta_1,\theta_2) = E[v(\theta_1,X(0))v(\theta_2,X(0))]\sigma^2$$

as $T \to \infty$.

4. Strong consistency

Let us now consider the limiting process $R_T(\theta)$ defined by (2.9). Any estimator $\hat{\theta}_T$ which minimizes

$$(4.1) \quad R_T(\theta) \equiv \int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt$$

$$-2\int_0^T [f(\theta,X(t))-f(\theta_0,X(t))]d\xi(t)$$

is called a process least squares estimator of θ.

Let μ_θ be the measure generated by the process X on C[0,T] when θ is the true parameter. From the general theory of diffusion processes, the Radon-Nikodym derivative of μ_θ with respect to μ_{θ_0} exists and is given by

$$(4.2) \quad \frac{d\mu_\theta}{d\mu_{\theta_0}} = \exp\left\{\int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}d\xi(t)\right.$$

$$\left. - \frac{1}{2}\int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt\right\} .$$

(cf. Gikhman and Skorokhod (1972), p.90). Hence

$$\log \frac{d\mu_\theta}{d\mu_{\theta_0}} = -\frac{1}{2} R_T(\theta)$$

which proves that the process least squares estimator $\hat{\theta}_T$ is the same as the maximum likelihood estimator $\tilde{\theta}_T$ of θ(cf. Basawa and Prakasa Rao (1979)) when the process X is observed over [0,T].

Let

$$(4.3) \qquad I_T(\theta) \equiv \int_0^T [f(\theta,X(t))-f(\theta_0,X(t))]^2 dt$$

and W^* be a standard Wiener process. Since the solution of the stochastic differential equation given in Section 2 is stationary ergodic by hypothesis, it follows that $I_T(\theta) \to \infty$ a.s. for $\theta \neq \theta_0$ by (A5) and the process $\{R_T(\theta)\}$ can be identified with the process $\{I_T(\theta) + 2W^*(T_T(\theta))\}$. Furthermore

$$(4.4) \qquad I_T(\theta) + 2W^*(T_T(\theta)) \to \infty \quad \text{a.s.}$$

as $T \to \infty$ for any $\theta = \theta_0$. Hence θ and θ_0 are pairwise consistent. Note that

$$(4.5) \qquad R_T(\theta) = I_T(\theta) + \sqrt{T}\, Z_T(\theta), \quad \theta \in \Theta, \quad T \geq 0$$

where $I_T(\theta)$ is defined by (4.3) and $Z_T(\theta)$ is given by (3.1). Let

$$(4.6) \qquad Z_T^*(\theta) = \sqrt{T}\, Z_T(\theta).$$

Then

$$(4.7) \qquad \frac{1}{T} I_T(\theta) \to I(\theta) \quad \text{a.s. as } T \to \infty \text{ by the ergodic theorem.}$$

In order to study the strong consistency of the estimator $\hat{\theta}_T$, we shall first obtain bounds on the modulus of coninuity of $I_T(\theta)$ and $Z_T^*(\theta)$.

<u>Lemma 4.1.</u> Under the assumptions (A1)-(A5),

$$|I_T(\theta)-I_T(\phi)| \leq C_1|\theta-\phi| \int_0^T J(X(t))(1+|X(t)|)dt \quad \text{a.s.}$$

where C_1 is a constant independent of T, θ and ϕ.

<u>Proof</u>. Note that

$$I_T(\theta)-I_T(\phi) = \int_0^T \{f(X(t),\theta)-f(X(t),\phi)\} \cdot$$

$$\cdot \{f(X(t),\phi)+f(X(t),\theta)-2f(X(t),\theta_0)\}dt$$

and therefore

$$|I_T(\theta)-I_T(\phi)| \leq |\theta-\phi| \int_0^T J(X(t))\cdot\{L(\theta)+L(\phi)+2L(\theta_0)\}\{1+|X(t)|\}dt$$

$$\leq C_1|\theta-\phi| \int_0^T J(X(t)\{1+|X(t)|\}dt .$$

Remark. Since $E[J^2(X(0)] < \infty$ and $E[X^2(0)] < \infty$, it follows that $E[J(X(0))X(0)] < \infty$ and hence by the ergodic theorem

$$\frac{1}{T}\int_0^T J(X(t))\{1+|X(t)|\}dt \xrightarrow{a.s.} E[J(X(0))\{1+|X(0)|\}] < \infty \quad \text{as } T \to \infty.$$

Therefore

(4.8) $$|I_T(\theta)-I_T(\phi)| \leq C^*T|\theta-\phi| \quad \text{a.s.}$$

as $T \to \infty$ for some constant $C^* > 0$. In view of (4.7) and Lemma 4.1, it follows that

(4.9) $$\frac{I_T(\theta)}{T} \xrightarrow{a.s.} I(\theta) \equiv E[f(\theta,X(0))-f(\theta_0,X(0))]^2$$

uniformly in $\theta \in \Theta$ as $T \to \infty$. But $I_T(\theta_0) = 0$ and $\underline{\lim}_{T\to\infty} \frac{I_T(\theta)}{T} > 0$ a.s. for $\theta \neq \theta_0$ by (A5). Hence, for any $\delta > 0$,

(4.10) $$\inf_{|\theta-\theta_0|\geq\delta} \frac{I_T(\theta)}{T} \xrightarrow{a.s.} \lambda \quad \text{as } T \to \infty$$

for some $\lambda > 0$ depending on δ.

Lemma 4.2. Under the assumptions (A1)-(A4), for any $T_0 > 0$ and any $\varepsilon > 0$,

(4.11) $$P(\sup_\theta \sup_{0\leq T\leq T_0} |Z_T^*(\theta)| > \varepsilon) \leq C_2 \frac{T_0}{\varepsilon^2}$$

for some constant $C_2 > 0$.

Proof. Let $h(\theta,x)$ and $g(\theta,x)$ be defined as in Section 3 and

$$h(\theta,x) = \sum_n a_n(x)e^{\pi i n\theta}, \quad \theta \in [-1,1].$$

Let

$$W_n^* = \int_0^T a_n(x(t))d\xi(t) .$$

Since $g(\theta,x)$ is a cubic polynomial in θ with coefficients in x, it is easy to check, by Kolmogorov's inequality, that

$$\text{(4.12)} \qquad \sup_\theta \sup_{0\le T\le T_0} \left| \int_0^T g(\theta,X(t))d\xi(t)\right| = O_p(T_0^{\frac{1}{2}})$$

using the fact that $|\theta| \le 1$. On the other hand, for any $\varepsilon > 0$,

$$\text{(4.13)} \qquad P(\sup_\theta \sup_{0\le T\le T_0} \left|\sum_n \{\int_0^T a_n(X(t))d\xi(t)\}e^{\pi i\theta}\right| > \varepsilon)$$

$$\le P(\sup_{0\le T\le T_0} \sum_n \left|\int_0^T a_n(X(t))d\xi(t)\right| > \varepsilon)$$

$$\le \sum_n P(\sup_{0\le T\le T_0} \left|\int_0^T a_n(X(t))d\xi(t)\right| > \varepsilon_n)$$

(where $\sum\varepsilon_n \le \varepsilon$)

$$\le \sum_n \frac{1}{\varepsilon_n^2} \mathrm{Var}(\int_0^{T_0} a_n(X(t))d\xi(t))$$

(by Kolmogorov's inequality for martingales)

$$\le \sum_n \frac{1}{\varepsilon_n^2} \int_0^{T_0} E(a_n(X(t)))^2 dt$$

$$= T_0 \sum_n \frac{\mu_n}{\varepsilon_n^2}$$

$$= \frac{T_0}{\varepsilon^2} (\sum\mu_n^{1/3})^3$$

when ε_n is chosen to be $\varepsilon\mu_n^{1/3} (\sum_n \mu_n^{1/3})^{-1}$. Note that $M \equiv \sum \mu_n^{1/3} < \infty$. Hence relations (4.12) and (4.13) together prove that

$$P(\sup_{\theta} \sup_{0 \le T \le T_0} |Z_T^*(\theta)| > \varepsilon) \le C_2 \frac{T_0}{\varepsilon^2} .$$

for some constant $C_2 > 0$ independent of T_0 and ε.

Lemma 4.3. For any $\gamma > 1/2$, there exists $H > 0$ such that

$$\text{(4.14)} \qquad \limsup_{T \to \infty} \sup_{\theta} \frac{|Z_T(\theta)|}{T^{1/2}(\log T)^{\gamma}} \le H \quad \text{a.s.}$$

Proof. Let

$$A_n = [\sup_{2^{n-1} < T \le 2^n} \sup_{\theta} |Z_T(\theta)| > H' \, 2^{n/2} n^{\gamma}], \quad n \ge 1.$$

Observe that Lemma 4.2 gives the inequality

$$P(A_n) = P[\sup_{0 < T \le 2^{n-1}} \sup_{\theta} |Z_T(\theta)| > H' 2^{n/2} n^{\gamma}]$$

(by stationarity of the process $X(t)$)

$$\le \frac{C\, 2^{n-1}}{H'^2 2^n n^{2\gamma}} = \frac{C}{2H'^2} \; \frac{1}{n^{2\gamma}} .$$

Hence $\sum_{n=1}^{\infty} P(A_n) < \infty$ which implies that $P(A_n$ occurs infinitely often$) = 0$ by Borel-Cantelli Lemma. Therefore $\sup_{\theta} |Z_T(\theta)| \le H' \, 2^{n/2} n^{\gamma}$ for all $2^{n-1} < T \le 2^n$ except for finitely many n with probability one and hence

$$\limsup_{T \to \infty} \sup_{\theta} |Z_T(\theta)| \le H \, T^{1/2} (\log T)^{\gamma} \quad \text{a.s.}$$

for suitable $H > 0$ depending on γ.

Theorem 4.1. Under the assumptions (A1)-(A5),

$$\hat{\theta}_T \to \theta_0 \quad \text{a.s. as } T \to \infty .$$

Proof. Note that

$$R_T(\theta) = I_T(\theta) + Z_T^*(\theta)$$

and $R_T(\theta_0) = 0$. Furthermore, for any $\delta > 0$, there exists $\lambda > 0$ depending on δ such that

$$\inf_{|\theta-\theta_0|\geq\delta} I_T(\theta) \geq T\lambda \quad \text{a.s.} \quad \text{as } T \to \infty$$

by (4.10) and with probability one, for any $\gamma > \frac{1}{2}$, there exists $H > 0$ depending on γ such that

$$\sup_{\theta} |Z_T^*(\theta)| \leq H\, T^{1/2}(\log T)^{\gamma} \quad \text{a.s.}$$

for sufficiently large T. Hence

$$\inf_{|\theta-\theta_0|\geq\delta} R_T(\theta) \geq \lambda^* T > 0 \quad \text{a.s.} \quad \text{as } T \to \infty.$$

for some $\lambda^* > 0$ depending on δ and γ. Since $\hat{\theta}_T$ minimizes $R_T(\theta)$ and $R_T(\theta_0) = 0$, it follows that $|\hat{\theta}_T-\theta_0| \leq \delta$ a.s. as $T \to \infty$. Hence $\hat{\theta}_T \to \theta_0$ a.s. as $T \to \infty$.

5. Asymptotic normality of the estimator

In addition to the conditions (A1)-(A5) assumed in Section 2, let us suppose that there exists a neighbourhood V_{θ_0} of θ_0 such that

(A6) $$|f_\theta^{(1)}(\theta,x)| \leq M(\theta)(1+|x|), \quad \theta \in V_{\theta_0}$$

and

$$\sup\{M(\theta): \theta \in V_{\theta_0}\} = M < \infty.$$

We shall now obtain the asymptotic distribution of $\hat{\theta}_T$ under the conditions (A1)-(A6). Since $\hat{\theta}_T$ is strongly consistent, $\hat{\theta}_T \in V_{\theta_0}$ with probability one for large T. Expanding $f(\theta,x)$ in a neighbourhood of θ_0, we have

$$f(\theta,x) = f(\theta_0,x) + (\theta-\theta_0)\tilde{f}(\tilde{\theta},x)$$

where $|\tilde{\theta}-\theta_0| \leq |\theta-\theta_0|$ and hence

(5.1) $$I_T(\theta) \equiv \int_0^T \{f(\theta,X(t))-f(\theta_0,X(t))\}^2 dt$$

$$= (\theta-\theta_0)^2 \int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt$$

$$+ (\theta-\theta_0)^2 \int_0^T [\{f_\theta^{(1)}(\tilde{\theta},X(t))\}^2-\{f_\theta^{(1)}(\theta_0,X(t))\}^2]dt.$$

Observe that

(5.2) $$|\{f_\theta^{(1)}(\tilde{\theta},x)\}^2-\{f_\theta^{(1)}(\theta_0,x)\}^2|$$

$$= |f_\theta^{(1)}(\tilde{\theta},x)-f_\theta^{(1)}(\theta_0,x)||f_\theta^{(1)}(\tilde{\theta},x)+f_\theta^{(1)}(\theta_0,x)|$$

$$\leq 2\, M|\tilde{\theta}-\theta_0|^\alpha\, c(x)(1+|x|)$$

by assumptions (A3) and (A6). Therefore

(5.3) $$|I_T(\theta) - (\theta-\theta_0)^2 \int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt|$$

$$\leq 2\, M|\theta-\theta_0|^{2+\alpha} \int_0^T c(X(t))(1+|X(t)|)dt.$$

Let us write $\theta-\theta_0 = T^{-1/2}\psi$. Then it follows that

(5.4) $$\sup_{|\psi|\leq A_T} |I_T(\theta)-\psi^2 T^{-1}\int_0^T\{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt| \leq M_1 A_T^{2+\alpha}T^{-1-\alpha}$$

for some constant $M_1 > 0$ by the erogodic theorem since

$$E(c(x(0))(1+|X(0)|) < \infty.$$

On the other hand, let

$$v_T(\psi,x) = T^{1/2}[f(\theta_0+\psi T^{-1/2},x)-f(\theta_0,x)-\psi T^{-1/2} f_\theta^{(1)}(\theta_0,x)]$$

for $|\psi| \leq A_T$. Then $v_T(\psi,x)$ is differentiable with respect to ψ and the derivative $v_T^{(1)}(\psi,x)$ satisfies

$$v_T^{(1)}(\psi,x)-v_T^{(1)}(\zeta,x) = f_\theta^{(1)}(\theta_0+\psi T^{-1/2},x)-f_\theta^{(1)}(\theta_0+\zeta T^{-1/2},x)$$

and hence

$$(5.5) \qquad |v_T^{(1)}(\psi,x)-v_T^{(1)}(\zeta,x)| \leq c(x)T^{-\alpha/2}|\psi-\zeta|^\alpha$$

by (A3) for all ψ, ζ in $[-A_T,A_T]$. It can be shown that there exists a polynomial in ψ with coefficients in x viz

$$(5.6) \qquad g_T(\psi,x) = v_T(A_T,x)P_1(\frac{\psi}{A_T}) + A_T v_T^{(1)}(A_T,x)P_2(\frac{\psi}{A_T})$$

$$+ v_T(-A_T,x)P_3(\frac{\psi}{A_T}) + A_T v_T^{(1)}(-A_T,x)P_4(\frac{\psi}{A_T})$$

on $[-A_T,A_T]$ such that

$$(5.7) \qquad g_T(A_T,x) = v_T(A_T,x), g_T(-A_T,x) = v_T(-A_T,x),$$

$$(5.8) \quad g_T^{(1)}(A_T,x) = v_T^{(1)}(A_T,x) \text{ and } g_T^{(1)}(-A_T,x) = v_T^{(1)}(-A_T,x)$$

where P_i, $1 \leq i \leq 4$ are polynomials in $\frac{\psi}{A_T}$ with constant coeffiecients. Observing that $v_T(0,x) = v_T^{(1)}(0,x) = 0$, it is easy to check that

$$(5.9) \qquad |g_T^{(1)}(A_T,x)| \leq c(x)A_T^\alpha T^{-\alpha/2} \quad ,$$

$$(5.10) \qquad |g_T^{(1)}(-A_T,x)| \leq c(x)A_T^\alpha T^{-\alpha/2} \quad ,$$

$$(5.11) \qquad |g_T(A_T,x)| \leq c(x)A_T^{1+\alpha}T^{-\alpha/2} \quad ,$$

and

$$(5.12) \qquad |g_T(-A_T,x)| \leq c(x)A_T^{1+\alpha}T^{-\alpha/2} \quad .$$

Furthermore there exists a constant $M_2 > 0$ independent of T such that

$$(5.13) \qquad |g_T^{(1)}(\psi,x)-g_T^{(1)}(\zeta,x)| \leq M_2 c(x)A_T^{\alpha-1}T^{-\alpha/2}|\psi-\zeta|$$

for all $\psi,\zeta \in [-A_T,A_T]$. But

$$A_T^{\alpha-1}|\psi-\zeta| \leq 2^{1-\alpha}|\psi-\zeta|^\alpha$$

since $|\psi-\zeta| \leq 2A_T$. Hence there exists a constant $M_3 > 0$ independent of T such that

$$(5.14) \qquad |g_T^{(1)}(\psi,x)-g_T^{(1)}(\zeta,x)| \leq M_3 c(x) T^{-\alpha/2}|\psi-\zeta|^\alpha$$

for all ψ, $\zeta\in[-A_T,A_T]$. Renormalizing, we get that

$$(5.15) \qquad |g_T^{(1)}(\psi^*,x)-g_T^{(1)}(\zeta^*,x)| \leq M_3 c(x) A_T^\alpha|\psi^*-\zeta^*|^\alpha T^{-\alpha/2}$$

for all $\psi^*,\zeta^* \in [-1,1]$. Let

$$(5.16) \qquad h_T(\psi^*,x) = v_T(\psi^*,x)-g_T(\psi^*,x).$$

Then there exists a constant $M_3^* > 0$ independent of T such that

$$(5.17) \qquad |h_T^{(1)}(\psi^*,x)-h_T^{(1)}(\zeta^*,x)| \leq M_3^* c(x) A_T^\alpha|\psi^*-\zeta^*|^\alpha T^{-\alpha/2}$$

for all $\psi^*,\zeta^* \in [-1,1]$ by relations (5.5) and (5.15). Now, applying Fourier series methods as in Lemma 4.2, it can be shown that for every $\varepsilon > 0$,

$$P(\sup_{|\psi^*|\leq 1} |\int_0^T v_T(\psi^*,X(t))d\xi(t)| > \varepsilon) \leq \frac{M_4 T}{\varepsilon^2} A_T^{2\alpha}T^{-\alpha}E[c^2(X(0)]$$

and hence

$$(5.18) \qquad P(\sup_{|\psi|\leq A_T} |\int_0^T \{f(\theta_0+\psi T^{-1/2},X(t))-f(\theta_0,X(t))$$

$$- \psi T^{-1/2} f_\theta^{(1)}(\theta_0,X(t))\}d\xi(t)| > \varepsilon)$$

$$\leq \frac{M_4}{\varepsilon^2} A_T^{2\alpha}T^{-\alpha}E[c^2(X(0)] .$$

Let us choose $A_T = \log T$. Since

$$\frac{1}{T}\int_0^T \{f_\theta^{(1)}(\theta_0,X(t))\}^2 dt \to I(\theta_0) = E[f_\theta^{(1)}(\theta_0,X(0))]^2 \quad \text{a.s.}$$

as $T \to \infty$ by the ergodic theorem and

$$\frac{1}{\sqrt{T}}\int_0^T f_\theta^{(1)}(\theta_0,X(t))d\xi(t) \xrightarrow{\mathcal{L}} N(0,\sigma^2 I(\theta_0)) \quad \text{as } T \to \infty$$

by the central limit theorem for stochastic integrals (cf. Basawa and Prakasa Rao (1979)), relations(5.4) and (5.18) imply that the asymptotic distribution of $\hat{\theta}_T$ which minimizes $R_T(\theta)$ given by (2.9) can be obtained from the process

$$(5.19) \qquad \psi^2 I(\theta_0) - 2\psi Z, \quad -\infty < \psi < \infty$$

where Z is normal with mean 0 and variance $\sigma^2 I(\theta_0)$. Since

$$\hat{\psi} = Z/I(\theta_0)$$

minimizes (5.16), it follows that

$$(5.20) \qquad T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} N(0,\sigma^2/I(\theta_0)).$$

This result is obtained under stronger conditions in Prakasa Rao (1979b) for the least squares estimator $\hat{\theta}_{n,T}$ defined at the beginning of Section 2.

Appendix

<u>Lemma 1</u> Suppose $\phi(u)$ is square integrable on [-1,1] and $\phi(\cdot)$ is Lipschitz of order α i.e., then exists c > 0 such that

$$(1) \qquad |\phi(u)-\phi(v)| \le c|u-v|^\alpha .$$

Let $\phi(u) = \sum_n a_n e^{\pi i n u}$. Then for any $0 < \gamma < \alpha$,

$$(2) \qquad \sum_n |a_n|^2 n^{2\gamma} \le K_1(\alpha,\gamma)c^2 .$$

Proof. It is easy to check that

$$(3)\qquad \int_{-1}^{1} |\phi(u+h)-\phi(u-h)|^2 du = 4 \sum_n |a_n|^2 \sin^2 \pi nh.$$

Since ϕ is Lipschitz satisfying (1), it follows that

$$(4)\qquad 4 \sum_n |a_n|^2 \sin^2 \pi nh \le 2^{2\alpha+1} c^2 h^{2\alpha}$$

for all $h \in [0,1]$. Let $h = 2^{-k}$ and $2^{k-2} < n \le 2^{k-1}$. It is clear that $\sin^2 \pi nh \ge \frac{1}{2}$ and relation (4) shows that

$$(5)\qquad \sum_{n=2^{k-2}+1}^{2^{k-1}} |a_n|^2 \le 2^{2\alpha} c^2 2^{-2k\alpha}$$

for any $k \ge 2$ and hence for any $0 < \gamma < \alpha$,

$$(6)\qquad \sum_{n=2^{k-2}+1}^{2^{k-1}} |a_n|^2 n^{2\gamma} \le 2^{2\alpha} c^2 2^{(2\gamma-2\alpha)k} .$$

Summing over all $k \ge 2$, we obtain that

$$(7)\qquad \sum_n |a_n|^2 n^{2\gamma} \le 2^{2\alpha} c^2 (1-2^{(2\gamma-2\alpha)})^{-1} .$$

Hence there exists a constant $K_1(\alpha,\gamma) > 0$ such that

$$(8)\qquad \sum_n |a_n|^2 n^{2\gamma} \le K_1(\alpha,\gamma) c^2$$

where c is the Lipschitzian constant given by (1).

Remark. A slight variation of the above result is due to Szasz (1922). The proof given above is the same as in Szasz (1922) and is given here for completeness.

Lemma 2. Suppose $h(u)$ is square integrable on $[-1,1]$ with $h(-1)=h(1)=0$ and $h'(\cdot)$ exists and is Lipschitzian of order α i.e., there exists $c > 0$ such that

(9) $$|h'(u)-h'(v)| \leq c|u-v|^{\alpha}.$$

Let $h(u) = \sum_n a_n e^{\pi i n u}$. Then, for any $0 < \gamma < \alpha$,

(10) $$\sum_n |a_n|^2 n^{2+2\gamma} \leq K_2(\alpha,\gamma)c^2$$

and

(11) $$\sum_n |a_n|^{2/3} \leq K_3(\alpha,\gamma)c^2 .$$

Proof. Since $h'(u) = \pi i \sum_n n a_n e^{\pi i n u}$, inequality (10) follows from Lemma 1. Observe that

$$\sum_n |a_n|^{2/3} \leq (\sum |a_n|^2 n^{2+2\gamma})^{1/3} (\sum n^{-(1+\gamma)})^{2/3}$$

$$\leq K_2(\alpha,\gamma)c^2 (\sum n^{-(1+\gamma)})^{2/3}$$

$$= K_3(\alpha,\gamma)c^2 .$$

Lemma 3. Let $h(\theta,x) = \sum_n a_n(x) e^{\pi i n \theta}$ and suppose there exists $\alpha > 0$ such that

$$|h_\theta^{(1)}(\theta,x) - h_\theta^{(1)}(\phi,x)| \leq c(x)|\theta-\phi|^{\alpha}$$

for all θ, ϕ in $[-1,1]$ where $f_\theta^{(1)}$ denotes the partial derivative of f with respect to θ. Let $\{X(t),\ t \in [0,T]\}$ be a stochastic process such that

$$E[h(\theta,X(t))]^2 < \infty$$

for every $t \in [0,T]$. Then, for any $\gamma < \alpha$, there exists a positive constant $K_4(\alpha,\gamma)$ such that

$$\sum_n \{\frac{1}{T}\int_0^T E[a_n^2(X(t))]dt\}^{1/3} \leq K_4(\alpha,\gamma)\{\frac{1}{T}\int_0^T E(c^2(X(t))dt\}^{1/3}.$$

Proof. By Lemma 2, it follows that

$$\sum_n |a_n(X(t))|^2 n^{2+2\gamma} \leq K_2(\alpha,\gamma)c^2(X(t)) \quad \text{a.s.}$$

for every $t \in [0,T]$. Hence

$$\sum_n E[a_n^2(X(t))]n^{2+2\gamma} \leq K_2(\alpha,\gamma)E[c^2(X(t))]$$

for all $t \in [0,T]$. Let

$$\mu_n = \frac{1}{T}\int_0^T E[a_n^2(X(t))]dt.$$

The inequality proved above gives the relation

$$\sum_n \mu_n n^{2+2\gamma} \leq K_2(\alpha,\gamma)\frac{1}{T}\int_0^T E[c^2(X(t))]dt$$

and hence

$$\begin{aligned}\sum_n \mu_n^{1/3} &\leq (\sum \mu_n n^{2+2\gamma})^{1/3}(\sum n^{-(1+\gamma)})^{2/3} \\ &\leq K_2^{1/3}(\alpha,\gamma)(\sum n^{-(1+\gamma)})^{2/3}\{\frac{1}{T}\int_0^T E[c^2(X(t))]dt\}^{1/3} \\ &\leq K_4(\alpha,\gamma)\{\frac{1}{T}\int_0^T E[c^2(X(t))]dt\}^{1/3}.\end{aligned}$$

Remark. Analgous argument proves that

$$\begin{aligned}\sum \mu_n^{1/2} &\leq (\sum \mu_n \cdot n^{2+2\gamma})^{1/2}(\sum n^{-2(1+\gamma)})^{1/2} \\ &< \infty .\end{aligned}$$

Acknowledgement. One of the authors (B.L.S.Prakasa Rao) thanks the Departments of Statistics and Mathematics of Purdue University for inviting him to spend the Summer 1979 which made the colloboration possible.

REFERENCES

[1] Basawa, I.V., Prakasa Rao, B.L.S. (1979). Statistical Inference for Stochastic Processes, Theory and Methods, Academic Press, London. (To appear).

[2] Dorogovcev, A. Ja. (1976). The consistency of an estimate of a parameter of a stochastic differential equaiton, Theory of Probability and Math. Statist. 10, 73-82.

[3] Gikhman, I.I. and Skorokhod, A.V. (1972). Stochastic Differential Equations, Springer-Verlag, Berlin.

[4] McKean, H. P. (1969). Stochastic Integrals, Academic Press, New York.

[5] Prakasa Rao, B.L.S. (1979a). The Bernstein-von Mises theorem for a class of diffusion processes, (Preprint) The University of Poona.

[6] Prakasa Rao, B.L.S. (1979b). Asymptotic theory for non-linear least squares estimators for diffusion processes, (Preprint), Indian Statistical Institute, New Delhi.

[7] Szász, O. (1922). Über den Kongvergenzexponent der Fourierschen Reihen, Münchener Sitzungskerichte, 135-150.

Department of Statistics
Purdue University
West Lafayette, IN 47907

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Mimeograph Series #79-13	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Asymptotic Theory for Process Least Squares Estimators for Diffusion Processes	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: Mime. Series #79-13
7. AUTHOR(s): B.L.S.Prakasa Rao and Herman Rubin	8. CONTRACT OR GRANT NUMBER(s): ONR-N00014-75-C-0455 NSF-MCS76-08316
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Purdue University, West Lafayette, Indiana 47907	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Washington, D.C.	12. REPORT DATE: July 1979
	13. NUMBER OF PAGES: 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

99173U

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stochastic Differential Equation; Diffusion

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Strong consistency and asymptotic normality of an estimator related to least squares estimator for parameters involved in nonlinear stochastic differential equations are investigated by studying families of stochastic integrals using Fourier analytic methods.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)